Constructions of MDS-Convolutional Codes

Roxana Smarandache, *Student Member, IEEE*,
Heide Gluesing-Luerssen, and
Joachim Rosenthal, *Senior Member, IEEE*

***Abstract*—Maximum-distance separable (MDS) convolutional codes are characterized through the property that the free distance attains the generalized Singleton bound. The existence of MDS convolutional codes was established by two of the authors by using methods from algebraic geometry. This correspondence provides an elementary construction of MDS convolutional codes for each rate k/n and each degree δ. The construction is based on a well-known connection between quasi-cyclic codes and convolutional codes.**

***Index Terms*—Convolutional codes, generalized Singleton bound, maximum-distance separable (MDS) convolutional codes.**

I. Introduction

The free distance of a rate k/n convolutional code of degree δ is always upper-bounded by the generalized Singleton bound

$$d_{\text{free}} \leq (n-k)(\lfloor \delta/k \rfloor + 1) + \delta + 1 \tag{1.1}$$

see [1]. We will provide an alternative proof of this result in the next section. If $\delta = 0$, i.e., in the case of block codes, (1.1) simply reduces to the well-known Singleton bound

$$d_{\text{free}} \leq n - k + 1 \tag{1.2}$$

cf. [2, Ch. 1, Theorem 11]. The authors of [1] showed the existence of rate k/n convolutional codes of degree δ whose free distance was equal to the generalized Singleton bound (1.1) and they called such codes maximum-distance separable (MDS) convolutional codes. The existence was established in [1] by techniques from algebraic geometry without giving an explicit construction. This correspondence is based on ideas from Justesen [3] and it provides an explicit construction of MDS convolutional codes for each rate k/n and each degree δ. The construction itself uses a well-known connection between quasi-cyclic codes and convolutional codes which has been worked out by several authors [3]–[6].

The correspondence is structured as follows. In the remainder of this section we introduce the basic notions which will be needed throughout the correspondence. In Section II, we give a new derivation of the generalized Singleton bound (1.1). The main new results will be given in Section III.

Let $\mathbb{F}$ be a finite field, $\mathbb{F}[D]$ the polynomial ring, and $\mathbb{F}(D)$ the field of rational functions. Let $G(D)$ be a $k \times n$ matrix over the polynomial ring $\mathbb{F}[D]$, with $\operatorname{rank} G(D) = k$. For the purpose of this correspondence, we define the rate k/n convolutional code generated by $G(D)$ as the set

$$\mathcal{C} = \{u(D)G(D) \in \mathbb{F}^n(D) | u(D) \in \mathbb{F}^k(D)\}$$

and say that $G(D)$ is a *generator matrix* for the convolutional code $\mathcal{C}$. If the generator matrices $G(D)$ and $G'(D)$ both generate the same convolutional code $\mathcal{C}$ then there exists a $k \times k$ invertible matrix $U(D)$ with $G'(D) = U(D)G(D)$ and we say $G(D)$ and $G'(D)$ are equivalent encoders.

Because of this, we can assume without loss of generality that the code $\mathcal{C}$ is presented by a *minimal basic encoder* $G(D)$. For this, let ν_i be the ith-row degree of $G(D)$, i.e., $\nu_i = \max_j \deg g_{ij}(D)$. In the literature [7], the indexes ν_i are also called the *constraint length for the* ith *input* of the matrix $G(D)$. Then one defines the following.

Definition 1.1: A polynomial generator matrix $G(D)$ is called *basic* if it has a polynomial right inverse. It is called *minimal* if $\sum_{i=1}^k \nu_i$ attains the minimal value among all generator matrices of $\mathcal{C}$.

A basic generator matrix is automatically *noncatastrophic*, this means finite-weight codewords can only be produced from finite-weight messages. If $G(D)$ is a minimal basic encoder one defines the *degree* [8] of $\mathcal{C}$ as the number $\delta := \sum_{i=1}^k \nu_i$. In the literature, the degree δ is sometimes also called the *total memory* [9] or the *overall constraint length* [7] or the *complexity* [10] of the minimal basic generator matrix $G(D)$, a number dependent only on $\mathcal{C}$. Among all these equivalent expressions we like the term degree best since it relates naturally to equal objects appearing in systems theory and algebraic geometry. The following remarks explain this notion.

Remark 1.2: It has been shown by Forney [11] that the set $\{\nu_1, \ldots, \nu_k\}$ of row degrees is the same for all minimal basic encoders of $\mathcal{C}$. Because of this reason, McEliece [8] calls these indexes the Forney indexes of the code $\mathcal{C}$. These indexes are also the same as the Kronecker indexes of the row module

$$\mathcal{M} = \{u(D)G(D) \in \mathbb{F}^n[D] | u(D) \in \mathbb{F}^k[D]\}$$

when $G(D)$ is a basic encoder. The Pontryagin dual of $\mathcal{M}$ defines a linear time-invariant behavior in the sense of Willems [12], [13], i.e., a linear system. Under this identification, the Kronecker indexes of $\mathcal{M}$ correspond to the observability indexes of the linear system [14]. The sum of the observability indexes is equal to the McMillan degree of the system. Finally, $\mathcal{M}$ defines in a natural way a quotient sheaf [15] over the projective line and, in this context, one refers to the indexes $\{\nu_1, \ldots, \nu_k\}$ as the Grothendieck indexes of the quotient sheaf and $\delta = \sum_{i=1}^k \nu_i$ as the degree of the quotient sheaf.

We feel that the degree is the single most important code parameter on the side of the transmission rate k/n. In the sequel, we will adopt the notation used by McEliece [8, p. 1082] and denote by (n, k, δ) a rate k/n convolutional code of degree δ.

For any n-component vector $v \in \mathbb{F}^n$, we define its weight $\operatorname{wt}(v)$ as the number of all its nonzero components. The weight $\operatorname{wt}(v(D))$ of a vector $v(D) \in \mathbb{F}^n(D)$ is then the sum of the weights of all its $\mathbb{F}^n$-coefficients. Finally, we define the *free distance* of the convolutional code $\mathcal{C} \subset \mathbb{F}^n(D)$ through

$$d_{\text{free}} = \min\{\operatorname{wt}(v(D)) | v(D) \in \mathcal{C},\, v(D) \neq 0\}. \tag{1.3}$$

It is an easy but crucial observation that in case we are given a basic encoder $G(D)$ the free distance can also be obtained as

$$d_{\text{free}} = \min\{\operatorname{wt}(v(D)) | v(D) \in \mathcal{M},\, v(D) \neq 0\}.$$

Manuscript received September 15, 1999; revised November 30, 2000. This work was supported in part by NSF under Grants DMS-96-10389 and DMS-00-72383. The work of R. Smarandache was supported by a fellowship from the Center of Applied Mathematics at the University of Notre Dame. J. Rosenthal was carrying out this work while he was a Guest Professor at EPFL in Switzerland. The material in this correspondence was presented in part at the 2000 IEEE International Symposium on Information Theory, Sorrento, Italy, June 25–30, 2000.

R. Smarandache and J. Rosenthal are with the Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: Smarandache.1@nd.edu; Rosenthal.1@nd.edu).

H. Gluesing-Luerssen is with the Department of Mathematics, Universität Oldenburg, Postfach 2503, D-26111 Oldenburg, Germany (e-mail: gluesing@mathematik.uni-oldenburg.de).

Communicated by R. M. Roth, Associate Editor for Coding Theory.

Publisher Item Identifier S 0018-9448(01)04751-4.

This follows simply from the fact that, if $G(D)$ has a polynomial right inverse, a nonpolynomial message $u(D)$ would result in a nonpolynomial codeword $v(D)$, which, of course, has infinite weight.

In the sequel, we wish to link the free distance to two types of distances known from the literature. Following the approach in [16], [7] we shall define the column distances d_j^c and the row distances d_j^r. In order to do so let us suppose

$$G(D) = G_0 + G_1 D + G_2 D^2 + \cdots + G_{\nu_k} D^{\nu_k}$$

is an encoder with row degrees $\nu_1 \leq \cdots \leq \nu_k$. Denote by (1.4) the *semi-infinite sliding generator matrix*, as shown at the bottom of the page. Then the convolutional code can be defined as

$$\mathcal{C} = \{(u_0, u_1, \ldots, u_\gamma, \ldots) \cdot G | u_j \in \mathbb{F}^k, \text{ for } j = 0, 1, \ldots\}.$$

Then the jth-*order column distance* d_j^c is defined as the minimum of the weights of the truncated codewords $v_{[0,j]} := (v_0, v_1, \ldots, v_j)$ resulting from an information sequence $u_{[0,j]} := (u_0, u_1, \ldots, u_j)$ with $u_0 \neq 0$. Precisely, if G_j^c denotes the $k(j+1) \times n(j+1)$ upper-left submatrix of the semi-infinite matrix G, then $d_j^c = \min_{u_0 \neq 0} \mathrm{wt}(u_{[0,j]} \cdot G_j^c)$. The quantity $d_{\nu_k}^c$ is called the *minimum distance* of the code and the tuple $d^{\boldsymbol{p}} = [d_0^c, d_1^c, \ldots, d_{\nu_k}^c]$ is called the *distance profile*. The limit $d_\infty^c = \lim_{j \to \infty} d_j^c$ exists and we have the relation

$$d_0^c \leq d_1^c \leq \cdots \leq d_\infty^c.$$

Then d_∞^c is the minimal distance computed over all finite or infinite codewords of $\mathcal{C}$. It is shown in [7] that $d_\infty^c = d_{\mathrm{free}}$.

The jth *row distance* d_j^r is defined as the minimum of the weights of all the finite codewords $v_{[0,j+\nu_k]} := (v_0, v_1, \ldots, v_{j+\nu_k})$ resulting from an information sequence $u_{[0,j]} := (u_0, u_1, \ldots, u_j) \neq 0$. Thus, if we denote by G_j^r the $k(j+1) \times n(j+\nu_k+1)$ upper-left submatrix of the semi-infinite matrix G, the jth-row distance is

$$d_j^r = \min_{u_{[0,j]} \neq 0} \mathrm{wt}(u_{[0,j]} \cdot G_j^r). \tag{1.5}$$

The limit $d_\infty^r = \lim_{j \to \infty} d_j^r$ exists and one has (see, e.g., [7]) for every encoder $G(D)$ the relation

$$d_0^c \leq d_1^c \leq \cdots \leq d_\infty^c = d_{\mathrm{free}} \leq d_\infty^r \leq \cdots \leq d_1^r \leq d_0^r. \tag{1.6}$$

In terms of state-space descriptions [17], [14] d_∞^r is equal to the minimal weight of a nonzero trajectory which starts from and returns to the all-zero state. d_∞^c is equal to the minimal weight of a nonzero trajectory which starts from and not necessarily returns to the all-zero state. Furthermore, if the generator matrix $G(D)$ is minimal basic, then $d_\infty^c = d_\infty^r = d_{\mathrm{free}}$ (see [17], [7] for details). It follows that for a basic encoder the minimal-weight codewords are generated by finite information sequences.

II. The Generalized Singleton Bound

It is certainly a most natural question to ask how large the distance of a rate k/n code of some bounded degree δ can be. McEliece [8] calls codes having the largest free distance among all (n, k, δ) codes *distance optimal*. Codes of degree $\delta = 0$ correspond to $[n, k]$ linear block codes and here we know that the distance cannot be larger than the Singleton bound $n - k + 1$. In [1], it was shown that the free distance can never be larger than the generalized Singleton bound (1.1) for an (n, k, δ)-code. In the sequel we will give a new derivation of this bound.

Once the row degrees $\nu_1, \ldots, \nu_k$ of the minimal basic encoder $G(D)$ are specified one has a natural upper bound on the free distance of a convolutional code. The following result was derived in [1].

Theorem 2.1: Let $\mathcal{C}$ be a rate k/n convolutional code generated by a minimal-basic encoding matrix $G(D)$. Let $\nu_1, \ldots, \nu_k$ be the row degrees of $G(D)$ and $\nu = \min\{\nu_1, \ldots, \nu_k\}$ denote the value of the smallest row degree. Finally, let ℓ be the number of indexes ν_i among the indexes $\nu_1, \ldots, \nu_k$ having the value ν. Then the free distance must satisfy

$$d_{\mathrm{free}} \leq n(\nu + 1) - \ell + 1. \tag{2.1}$$

The proof given in [1] was based on the polynomial generator matrix $G(D)$. In the sequel, we provide a proof by means of the sliding matrix G introduced in (1.4).

Proof: Without loss of generality, we may assume

$$\nu = \nu_1 \leq \cdots \leq \nu_k.$$

Let G be the infinite sliding generator matrix associated to $G(D)$ as in (1.4). We will show that the bound (2.1) is actually a bound on the 0th row distance d_0^r defined in (1.5),; in other words, we will show that $d_0^r \leq n(\nu + 1) - \ell + 1$. From this, the claim follows using (1.6). To prove the bound on d_0^r, we only need to look at the first block-row of the sliding matrix G denoted by

$$G_0^r = [G_0\ G_1\ \cdots\ G_{\nu_1}\ G_{\nu_1+1}\ \cdots\ G_{\nu_k}].$$

For all $j = 0, \ldots, \nu_k$ let G_j' denote the $\ell \times n$ matrix formed by the first ℓ rows of the matrix G_j. All matrices $G_{\nu_1+1}', \ldots, G_{\nu_k}'$ are zero. Hence the minimum distance d_0^r of the $[n(\nu_k + 1), k]$ block code generated by $[G_0\ \ G_1\ \ \cdots\ \ G_{\nu_1}\ \ G_{\nu_1+1}\ \ \cdots\ \ G_{\nu_k}]$ is smaller than the minimum distance of the $[n(\nu + 1), \ell]$ block code generated by ${G'}_0^r := [G_0'\ G_1'\ \cdots\ G_{\nu_1}']$, which is upper-bounded by the Singleton bound $n(\nu + 1) - \ell + 1$. Therefore, we obtain the desired bound on d_0^r and hence on the free distance

$$d_{\mathrm{free}} = d_\infty^r \leq \cdots \leq d_2^r \leq d_1^r \leq d_0^r \leq n(\nu + 1) - \ell + 1. \qquad \square$$

Remark 2.2: It was pointed out to the authors by a referee that Theorem 2.1 can also be derived from [8, Theorem 4.4] and [8, Corollary 4.3].

In the case of a block code, i.e., when $\nu = 0$ and $\ell = k$, the upper bound in (2.1) is identical to the Singleton bound (1.2).

It is easy to see that for given n, k, and δ, the upper bound (2.1) is maximized if and only if ν is as big as possible while ℓ is as small as possible, which results in

$$\begin{aligned} \nu = \lfloor \delta/k \rfloor = \nu_1 = \cdots = \nu_\ell < \nu_{\ell+1} = \\ \cdots = \nu_k = \lfloor \delta/k \rfloor + 1 = \nu + 1. \end{aligned} \tag{2.2}$$

We will call the above set of indexes the *generic set of row degrees* as they are sometimes referred to in the systems literature.

Remark 2.3: McEliece [8, p. 1083] calls a code $\mathcal{C}$ having the generic set of row degrees *compact*. In systems theory, the set of row degrees $\nu_1, \ldots, \nu_k$ corresponds to the *observability indexes* of the associated (Pontryagin dual) linear system. (Compare with Remark 1.2 and [14]). It is known that the set of all linear systems having a fixed input number k, a fixed output number $n - k$, and a fixed McMillan degree δ has in a

$$G = \begin{bmatrix} G_0 & G_1 & \cdots & G_{\nu_1} & G_{\nu_1+1} & \cdots & G_{\nu_k} & & \\ & G_0 & G_1 & \cdots & G_{\nu_1} & G_{\nu_1+1} & \cdots & G_{\nu_k} & \\ & & \ddots & \ddots & & \ddots & \ddots & & \ddots \end{bmatrix}. \tag{1.4}$$

natural way the structure of a smooth projective variety [15]. The subset of systems having the generic set of row degrees forms a Zariski open subset of this variety, i.e., a generic set in the sense of algebraic geometry. This explains why systems theorists call the indexes appearing in (2.2) the generic set of row degrees.

Specializing the above result to the generic set of row degrees we get the following upper bound in terms of the degree δ.

Theorem 2.4: For every base field $\mathbb{F}$ and every rate k/n convolutional code $\mathcal{C}$ of degree δ, the free distance is bounded by

$$d_{\text{free}} \leq (n-k)(\lfloor \delta/k \rfloor + 1) + \delta + 1. \tag{2.3}$$

The main result of [1] states.

Theorem 2.5: For any positive integers $k < n,\ \delta$ and for any prime p there exists a rate k/n convolutional code $\mathcal{C}$ of degree δ over a sufficiently big field of characteristic p, whose free distance is equal to the upper bound (2.3).

Based on Theorems 2.4 and 2.5 we introduce the following notions.

Definition 2.6: The upper bound (2.3) is called the *generalized Singleton bound*. A rate k/n code of degree δ whose free distance achieves the generalized Singleton bound is called an *MDS convolutional code*.

The proof of Theorem 2.5 given in [1] is nonconstructive and it makes use of algebraic geometry. For some special set of rates k/n and degree δ, e.g., $k = 1$ [18] or $k = \delta - 1$ [19], constructions which lead to MDS convolutional codes can be found in the literature. We are, however, not aware of a construction in the general case.

The algebraic conditions used in [1] to describe the set of MDS convolutional codes were very involved and we do not know of a simple algebraic criterion in general. For small parameters $k,\ n$ and δ it is, however, often easy to decide if a particular code is MDS. The following example illustrates this.

Example 2.7: Consider the rate $2/3$ convolutional code over the base field $\mathbb{F}_3$ defined through the encoding matrix

$$G(D) = \begin{bmatrix} 1 & 1 & 1 \\ D+1 & D & 2D+2 \end{bmatrix}.$$

Here the row degrees are $\nu = \nu_1 = 0$ and $\nu_2 = 1$, $\ell = 1$ and the total degree is $\delta = 1$. $\nu_1,\ \nu_2$ form a generic set of row degrees and the upper bounds in (2.1) and (2.3) are in this case both equal to 3.

It follows that $G(D)$ is an MDS convolutional code if the free distance of this code is equal to 3. One verifies that the 0th column distance $d_0^c = 2$ and the first column distance is $d_1^c = 3$, the maximal possible.

It follows from Theorem 2.1 that MDS convolutional codes necessarily have the generic set of row degrees as in (2.2). It is worth mentioning that within the class of all rate k/n codes with fixed degree δ, the distribution (2.2) of the row degrees leads to the smallest possible memory.

The set of convolutional codes of rate k/n and degree δ is subdivided into codes whose encoding matrices $G(D)$ have a fixed set of row degrees $\nu_1,\ \ldots,\ \nu_k$ with $\delta = \sum_{i=1}^{k} \nu_i$. In Theorem 2.1, we gave an upper bound for the free distance for a code whose row degrees are not necessarily the generic set of indexes. It is an open question if there always exist convolutional codes having given row degrees $\nu = \nu_1 \leq \cdots \leq \nu_k$ and free distance equal to the right-hand side of (2.1).

We conclude the section with a simple theorem that tells us how to obtain MDS convolutional codes of rate k'/n from MDS codes of rate k/n where $k' < k$.

Theorem 2.8: Let $\mathcal{C}$ be a convolutional code of rate k/n generated by the minimal-basic encoding matrix $G(D) \in \mathbb{F}[D]^{k \times n}$ with row indexes

$$\nu = \nu_1 = \cdots = \nu_\ell < \nu_{\ell+1} \leq \cdots \leq \nu_k, \qquad \text{where } \ell < k.$$

Let $\overline{G}(D) \in \mathbb{F}[D]^{(k-1)\times n}$ be the matrix obtained from $G(D)$ by omitting any of the last $k - \ell$ last rows of $G(D)$. If the free distance of $\mathcal{C}$ achieves the upper bound (2.1), then the same is true for the code $\overline{\mathcal{C}}$ generated by the encoder $\overline{G}$. In particular, if $\mathcal{C}$ is an MDS code, then so is $\overline{\mathcal{C}}$.

Proof: First note that noncatastrophicity as well as the full-rank conditions carry over to the matrix $\overline{G}$. Moreover, the codes $\mathcal{C}$ and $\overline{\mathcal{C}}$ both have the same minimal row degree ν and the same number ℓ of rows having this degree ν. Therefore, the upper bound (2.1) has the same value for both codes and the theorem follows from the inclusion $\overline{\mathcal{C}} \subseteq \mathcal{C}$. □

III. A Construction of Rate k/n MDS Convolutional Codes

In this section, we will provide a concrete construction of an $(n,\ k,\ \delta)$ MDS convolutional code for each degree δ and each rate k/n. The underlying idea here follows the lines of [3], [5] which is an instance of the relationship between quasi-cyclic block codes and convolutional codes. We will not go into the details of this connection, rather refer the reader to [3], [4], [6].

As defined in [3], [5], a convolutional code is said to be *generated by a polynomial*

$$g(D) = g_0(D^n) + g_1(D^n)D + \cdots + g_{n-1}(D^n)D^{n-1} \tag{3.1}$$

if it has a polynomial encoder of the form (3.2) shown in at the bottom of the page. It is immediate that $\operatorname{rank} G(0) = k$ if $g(0) = g_0(0) \neq 0$.

The code

$$\mathcal{C} = \{(u_0(D),\ \ldots,\ u_{k-1}(D)) \cdot G(D) \mid (u_0(D),\ \ldots,\ u_{k-1}(D)) \in \mathbb{F}^k[D]\}$$

is isomorphic to

$$\{(u_0(D^n) + u_1(D^n)D + \cdots + u_{k-1}(D^n)D^{k-1}) \cdot g(D)\} \tag{3.3}$$

the isomorphism is simply multiplexing and, therefore, weight-preserving. We will not use the description (3.3) but rather the encoder matrix in (3.2).

The following theorem will lead us to the construction of MDS convolutional codes. Recall that two elements $a,\ b \in \mathbb{F}$ are called n-equivalent if $a^n = b^n$.

Theorem 3.1 [3, Theorem 3]: Let p be a prime and $r \in \mathbb{N}$. Let $g(D) \in \mathbb{F}_{p^r}[D]$ generate a cyclic code over $\mathbb{F}_{p^r}$ of length N relatively prime to p and of distance d_g. Let n be any positive divisor of N and

$$G(D) = \begin{bmatrix} g_0(D) & g_1(D) & g_2(D) & \cdots & \cdots & \cdots & g_{n-1}(D) \\ Dg_{n-1}(D) & g_0(D) & g_1(D) & \cdots & \cdots & \cdots & g_{n-2}(D) \\ Dg_{n-2}(D) & Dg_{n-1}(D) & g_0(D) & \cdots & \cdots & \cdots & g_{n-3}(D) \\ \vdots & \vdots & \ddots & \ddots & & & \vdots \\ Dg_{n-k+1}(D) & Dg_{n-k+2}(D) & \cdots & Dg_{n-1}(D) & g_0(D) & \cdots & g_{n-k}(D) \end{bmatrix}. \tag{3.2}$$

$k < n$. If $g(D)$ has at most $n - k$ roots in each n-equivalence class, then the generator matrix $G(D)$ defined in (3.2) is basic minimal and describes a k/n convolutional code of free distance $d_{\text{free}} \geq d_g$.

Now we are ready to construct MDS codes of any rate k/n and any degree δ. The idea is as follows. We will construct a polynomial $g(D) \in \mathbb{F}_{p^r}[D]$ of degree $N - K$ which generates a rate $[N, K]$ Reed–Solomon block code whose distance is equal to the Singleton bound $N - K + 1$. The parameters N and K will be chosen such that $n|N$ and $d_g = (n-k)(\lfloor \delta/k \rfloor + 1) + \delta + 1$, which is the MDS bound for the given parameters n, k, and δ (see (2.3)). The polynomial $g(D)$ will satisfy the conditions of Theorem 3.1, thus we obtain the desired MDS convolutional code.

To accomplish this the following technical lemma will be needed.

Lemma 3.2: Let p be a prime and k, n, δ fixed positive integers such that p and n are relatively prime and $k < n$. Then there exist positive integers r and a

$$a \geq \lfloor \delta/k \rfloor + 1 + \delta/(n-k) \tag{3.4}$$

solving the Diophantine equation

$$an = p^r - 1. \tag{3.5}$$

Proof: Consider the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^*$ which has order $\phi(n)$. Since $(p, n) = 1$ we know that $p^{i\phi(n)} \equiv 1 \bmod n$ for all $i \geq 1$. In particular, $p^{i \cdot \phi(n)} - 1$ is divisible by n. Choose i such that (3.4) is satisfied for

$$a := \frac{p^{i \cdot \phi(n)} - 1}{n}. \qquad \square$$

In the sequel, assume that a, r is a solution of (3.5) satisfying the inequality (3.4). Let $N = an$ and let $K = N - (n-k)(\lfloor \delta/k \rfloor + 1) - \delta$. It is easily seen that $0 < K < N$. Let $\alpha \in \mathbb{F}_{p^r}$ be a primitive element of $\mathbb{F}_{p^r}$ and define

$$g(D) = (D - \alpha^0)(D - \alpha^1) \cdots (D - \alpha^{N-K-1}) \in \mathbb{F}_{p^r}[D]. \tag{3.6}$$

The polynomial $g(D)$ defines a rate $[N, K]$ Reed–Solomon block code with distance

$$d_g = N - K + 1 = (n-k)(\lfloor \delta/k \rfloor + 1) + \delta + 1$$

as desired.

Theorem 3.3: Let p, n, k and δ be integers with $k < n$ and n not divisible by p. Then there exists an MDS convolutional code of rate k/n and degree δ over some suitably big field of characteristic p. Indeed, the generator matrix $G(D)$ in (3.2) induced by the polynomial $g(D)$ given in (3.6) defines an MDS convolutional code of rate k/n and degree δ over $\mathbb{F}_{p^r}$.

Proof: First we show that the generator matrix $G(D)$ is of degree δ. In order to do so, we calculate the degrees of the polynomials $g_i(D)$ in the expansion (3.1) of $g(D)$. First note that

$$\deg g(D) = N - K = n\nu + n - \ell$$

where $\nu = \lfloor \delta/k \rfloor$ and $\ell = k(\lfloor \delta/k \rfloor + 1) - \delta > 0$. Since $g(D)$ defines a Reed–Solomon block code it follows that all its coefficient are nonzero and one obtains

$$\begin{aligned} &\deg g_i(D) = \nu, && \text{for } i = 0, \ldots, n - \ell \\ &\deg g_i(D) = \nu - 1, && \text{for } i = n - \ell + 1, \ldots, n-1. \end{aligned}$$

This implies that the row degrees of $G(D)$ are indeed as in (2.2) and that $G(D)$ is minimal. Thus, the degree of the code generated by $G(D)$ is simply given by the sum of the row degrees, which is in fact

$$\ell\nu + (k - \ell)(\nu + 1) = k(\lfloor \delta/k \rfloor + 1) - \ell = \delta.$$

Observe also that $\operatorname{rank} G(0) = k$.

Next we prove that g satisfies the root condition given in Theorem 3.1. To do so, observe that the n-equivalence class of α^s, where $0 \leq s \leq a - 1$, consists of

$$\alpha^s, \alpha^{s+a}, \alpha^{s+2a}, \ldots, \alpha^{s+a(n-k-1)}, \alpha^{s+a(n-k)}, \ldots.$$

The form of $g(D)$ in (3.6) shows that each such n-equivalence class contains at most $n - k$ roots of $g(D)$ if $N - K \leq (n-k)a$. This is indeed guaranteed by construction of a in (3.4)

$$a \geq \lfloor \delta/k \rfloor + 1 + \frac{\delta}{n-k} = \frac{N-K}{n-k}. \tag{3.7}$$

Now Theorem 3.1 implies that the encoder $G(D)$ given in (3.2) is minimal-basic and generates an MDS code with the given parameters n, k, and δ. $\square$

Remark 3.4: The above proof is quite similar to the proof of [3, Theorem 4]. Actually, Justesen's Theorem 4 can be considered a special case of the above, namely, the case when $K = ka$. In the above construction, we have more generally $K \geq ka$, see (3.7). The case $K = ka$ can occur only if $(n-k)|\delta$, which we did not require.

It is interesting to study the constructed convolutional code via the semi-infinite sliding generator matrix as introduced in (1.4). To do so we expand the generator polynomial $g(D)$ in terms of its coefficients

$$g(D) = c_0 + c_1 D + \cdots + c_{N-K} D^{N-K}.$$

The $[N, K]$ Reed–Solomon block code generated by $g(D)$ has a generator matrix of the form

$$\mathcal{G} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{N-K} & & & & \\ & c_0 & c_1 & \cdots & c_{N-K} & & & \\ & & \ddots & \ddots & & & \ddots & \\ & & & c_0 & & c_1 & \cdots & c_{N-K} \end{bmatrix}. \tag{3.8}$$

A direct calculation now shows that the first k rows of the matrix $\mathcal{G}$ appear as the upper-left corner of the matrix G in (1.4), where, again, the matrix $G(D)$ is as in (3.2). Thereafter, rows $jn + 1, \ldots, jn + k$ of $\mathcal{G}$ correspond to rows $jk + 1, \ldots, (j+1)k$ of G expressing the polynomial description (3.3). If $\mathcal{G}$ was an infinite sliding-block matrix it would trivially follow that the convolutional code $G(D)$ has free distance $d_{\text{free}} \geq N - K + 1$. Theorem 3.1 of Justesen and in particular the "weight retaining property" as studied by Massey, Costello, and Justesen [5] guarantee that the distance estimate holds for the semi-infinite sliding generator matrix G.

Remark 3.5: We formulated Theorem 3.3 with a prescribed characteristic p of the field over which we construct the MDS convolutional code. If one is interested in the smallest possible field where this construction works, regardless of characteristic, one should, of course, choose a to be the smallest integer such that $a \geq \lfloor \delta/k \rfloor + 1 + \delta/(n-k)$ and $an + 1$ is a prime power. In any case, it follows immediately from (3.4) and (3.5) that the field size is the smallest possible prime power q for which

$$n|(q-1) \quad \text{and} \quad q \geq \delta \frac{n^2}{k(n-k)} + 2. \tag{3.9}$$

We close this section with a few examples.

Example 3.6: Suppose we want to construct a $(3, 2, 5)$ MDS convolutional code. The MDS bound is in this case 9 and from (3.9) we need the smallest prime power p^r bigger than 24, such that $p^r - 1$ is divisible by 3. The smallest possible field is $\mathbb{F}_{5^2}$ and we will need a rate $[24, 16]$ Reed–Solomon code for the construction.

$$G(D) = \begin{bmatrix} \alpha^{28} + \alpha^{35}D + \alpha^{57}D^2 & 1 + \alpha^6 D + \alpha^{42}D^2 & \alpha^8 + \alpha^{26}D + D^2 \\ \alpha^8 D + \alpha^{26}D^2 + D^3 & \alpha^{28} + \alpha^{35}D + \alpha^{57}D^2 & 1 + \alpha^6 D + \alpha^{42}D^2 \end{bmatrix}.$$

If we want however an MDS code in characteristic 2, the smallest field is $\mathbb{F}_{2^6}$, and we need a rate $[63, 55]$ Reed–Solomon code. Using, e.g., MAPLE, one calculates

$$\begin{aligned} g(D) &= \prod_{i=0}^{7}(D - \alpha^i) \\ &= D^8 + \alpha^{42}D^7 + \alpha^{57}D^6 + \alpha^{26}D^5 + \alpha^6 D^4 + \alpha^{35}D^3 \\ &\quad + \alpha^8 D^2 + D + \alpha^{28} \\ &= (\alpha^{28} + \alpha^{35}D^3 + \alpha^{57}D^6) + D(1 + \alpha^6 D^3 + \alpha^{42}D^6) \\ &\quad + D^2(\alpha^8 + \alpha^{26}D^3 + D^6) \end{aligned}$$

where α is a primitive of $\mathbb{F}_{2^6}$. Hence, an encoder for a $(3, 2, 5)$ MDS convolutional code is given by the equation at the top of the page.

Example 3.7: Another example that we give is a $(5, 2, 12)$ MDS convolutional code. The MDS bound is $5(6+1) - 2 + 1 = 34$ and, as before, we will need the smallest prime power p^r bigger than 55, such that $p^r - 1$ is divisible by 5. The smallest possible field is $\mathbb{F}_{61}$ and we need a $[60, 27]$ Reed–Solomon code for the construction.

If we want to have the construction over a field of characteristic 2 we will have to take $a = 51$ in (3.5) which makes $N = q - 1 = 2^8 - 1 = 255$. The Reed–Solomon code that we use has parameters $N = 255$ and $K = 222$.

IV. Conclusion

In this correspondence, we constructed MDS convolutional codes for each rate k/n and for each code of degree δ. The construction was based on the construction of a large Reed–Solomon block code and because of this the obtained convolutional code is closely related to this Reed–Solomon code. The correspondence raises several follow-up questions. Is it possible to come up with an independent construction which does not require the relative primeness of the characteristic p and the length n of the code, and/or which does not need such large field sizes? Is it possible to carry through some subfield constructions and is it possible to come up with an algebraic decoding algorithm? Finally, it would be interesting to understand MDS convolutional codes from the point of view of state dynamics. Some answers in these directions were given in [17], [20] but more research is needed.

References

[1] J. Rosenthal and R. Smarandache, "Maximum distance separable convolutional codes," *Appl. Algebra Engrg. Comm. Comput.*, vol. 10, no. 1, pp. 15–32, 1999.

[2] F. J. MacWilliams and N. J. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North Holland, 1977.

[3] J. Justesen, "New convolutional code constructions and a class of asymptotically good time-varying codes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 220–225, Mar. 1973.

[4] Y. Levy and D. J. Costello, Jr., "An algebraic approach to constructing convolutional codes from quasicyclic codes," *DIMACS Ser. Discr. Math. Theor. Comput. Sci.*, vol. 14, pp. 189–198, 1993.

[5] J. L. Massey, D. J. Costello, Jr., and J. Justesen, "Polynomial weights and code constructions," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 101–110, Jan. 1973.

[6] R. M. Tanner, "Convolutional codes from quasicyclic codes: A link between the theories of block and convolutional codes," Univ. Calif., Santa Cruz, Tech. Rep. UCSC-CRL-87-21, Nov. 1987.

[7] R. Johannesson and K. S. Zigangirov, *Fundamentals of Convolutional Coding*. Piscataway, NJ: IEEE Press, 1999.

[8] R. J. McEliece, "The algebraic theory of convolutional codes," in *Handbook of Coding Theory*, V. Pless and W. Huffman, Eds. Amsterdam, The Netherlands: Elsevier Science, 1998, vol. 1, pp. 1065–1138.

[9] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[10] P. Piret, *Convolutional Codes, an Algebraic Approach*. Cambridge, MA: MIT Press, 1988.

[11] G. D. Forney, "Minimal bases of rational vector spaces, with applications to multivariable linear systems," *SIAM J. Contr.*, vol. 13, no. 3, pp. 493–520, 1975.

[12] J. Rosenthal, J. M. Schumacher, and E. V. York, "On behaviors and convolutional codes," *IEEE Trans. Inform. Theory*, pt. 1, vol. 42, pp. 1881–1891, Sept. 1996.

[13] J. C. Willems, "Paradigms and puzzles in the theory of dynamical systems," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 259–294, Mar. 1991.

[14] J. Rosenthal, "Connections between linear systems and convolutional codes," in *Codes, Systems and Graphical Models*, B. Marcus and J. Rosenthal, Eds. Berlin, Germany: Springer-Verlag, 2000, vol. 123, pp. 39–66.

[15] M. S. Ravi and J. Rosenthal, "A smooth compactification of the space of transfer functions with fixed McMillan degree," *Acta Appl. Math*, vol. 34, pp. 329–352, 1994.

[16] R. Johannesson and K. Zigangirov, "Distances and distance bounds for convolutional codes—An overview," in *Topics in Coding Theory. In honor of L. H. Zetterberg (Lecture Notes in Control and Information Sciences)*. Berlin, Germany: Springer Verlag, 1989, vol. 128, pp. 109–136.

[17] J. Rosenthal and E. V. York, "BCH convolutional codes," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1833–1844, Sept. 1999.

[18] J. Justesen, "An algebraic construction of rate $1/\nu$ convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 577–580, Jan. 1975.

[19] G. Lauer, "Some optimal partial-unit-memory codes," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 240–243, Mar. 1979.

[20] R. Smarandache, "Unit memory convolutional codes with maximum distance," in *Codes, Systems and Graphical Models*, B. Marcus and J. Rosenthal, Eds. Berlin, Germany: Springer-Verlag, 2000, vol. 123, pp. 381–396.